Machine Learning Based Clustering and Modeling for 6G UAV-to-Ground Communication Channels

Zhaolei Zhang, Yu Liu, *Member, IEEE,* Cheng-Xiang Wang, *Fellow, IEEE,* Hengtai Chang, *Member, IEEE,* Ji Bian, *Member, IEEE*, Jingfan Zhang

***Abstract*—Towards the sixth-generation (6G) wireless communication, unmanned aerial vehicles (UAVs) have been regarded as an indispensable part due to its flexible deployment, wide coverage, and high mobility. This also creates challenges for channel research. Scatterers are normally present in the structure of clusters during UAV communication, and cluster-based channel modeling is significant. In this paper, the variational Bayesian-Gaussian mixture model (VB-GMM) algorithm is proposed for clustering, which takes into account the time-space properties. Cluster tracking is implemented using the multipath component distance (MCD) algorithm. Intra- and inter-cluster characterization, such as the number of clusters, cluster power distribution, angular/delay offset, and angular/delay spreads, are well studied. Moreover, cluster lifetime and birth-death (B-D) properties are extracted and analyzed. Based on these cluster characteristics acquired by machine learning (ML) method, a novel UAV-to-ground communication channel model is proposed, and a four-state Markov chain is also introduced to portray the evolution of clusters. Simulation results match well with channel measurements, which verifies the practicality of the proposed model. This paper can give theoretical and technical support for the design and evaluation of UAV-to-ground communication systems.**

***Index Terms*—6G, UAV communications, cluster characterization, channel modeling, machine learning.**

I. Introduction

As an essential part of the future sixth-generation (6G) wireless communications and air-space-ground-sea integrated network, unmanned aerial vehicle (UAV) technologies are already attracting much attention [1]–[3]. Owing to its advantages of fast deployment, wide coverage, and high mobility, UAV communications can be widely used in urban transportation, target detection, and emergency rescue. Moreover, it is commonly recognized that the development, verification, and evaluation of 6G UAV-to-ground communication systems rely heavily on precise channel models [4], [5]. Therefore, it is vital to investigate the UAV-to-ground communication channels.

In radio propagation, the received multipath signals often exhibit the cluster structure [6]. The cluster is a set of multipaths in the wireless channel with similar characteristic attributes, e.g., angle of arrival (AoA), angle of departure (AoD), and delay. For UAV communication, channel clustering means grouping the channel multipath based on above similar characteristic parameters. Moreover, the prevalent approach based on the geometry-based stochastic models (GBSMs) also adopts the cluster-based methodology and has shown promising performance [7]–[12]. Some standardized documents also use the cluster-based structure, such as IMT-2020 [13], 5GCM-SIG [14], and 3GPP TR 38.901 [15]. Cluster-based channel models can reduce algorithm complexity while maintaining validity [16]. However, for UAV communications, the scatterer distribution and fast time-varying characteristics illuminate the complex structure of clusters, which creates difficulties for cluster-based channel research. Therefore, effective clustering and relevant analysis are essential.

Currently, wireless channel data present diverse and massive characteristics. Compared to conventional methods, there is no doubt that the machine learning (ML) method for clustering research is a powerful tool [17]. So far, several algorithms have been applied to channel clustering in different scenarios. The earliest clustering method was based on the Saleh Valenzuela (SV) model, but this model was oversimplified and did not reflect the channel characteristics comprehensively [18], [19]. In wireless communications, the K-means algorithm has been widely used, and the K-power-means (KPM) algorithm can be obtained by considering the power of multipath components (MPCs) to enhance their performance. The KPM algorithm has found widespread application in clustering for various channels, including indoor [20], outdoor-to-indoor (O2I) [21], and high-speed railway (HSR) [22] environments. It belongs to the category of hard clustering methods based on distance-based classification. A soft clustering method, i.e., fuzzy C-means (FCM), was employed for clustering analysis in the staircase environment at 60 GHz [23]. It is well known that MPCs of the same cluster can be modeled as a decaying power delay profile (PDP). Therefore, shape-based cluster identification can be used. The kurtosis measurement method was used to identify clusters in channel impulse responses (CIRs), and it applied a region competition approach to classify MPCs into different clusters [24]. Unlike distance-based clustering methods such as KPM and FCM, shape-based clustering

This work was supported by the National Natural Science Foundation of China (NSFC) under Grants 62001269, 62301365, and 62101311, the State Key Laboratory of Rail Traffic Control and Safety (Contract No. RCS2022K009), Beijing Jiaotong University, and the Future Plan Program for Young Scholars of Shandong University. *(Corresponding authors: Yu Liu.)*

Zhaolei Zhang, Yu Liu, and Jingfan Zhang are with School of Integrated Circuits, Shandong University, Jinan, 250101, China (e-mail: zhaoleizhang@mail.sdu.edu.cn; yuliu@sdu.edu.cn; 202132424@mail.sdu.edu.cn).

Cheng-Xiang Wang is with the National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University, Nanjing, 210096, China, and also with Purple Mountain Laboratories, Nanjing, 211111, China (e-mail: chxwang@seu.edu.cn).

Hengtai Chang is with with Purple Mountain Laboratories, Nanjing, 211111, China, and also with School of Information Science and Engineering, Southeast University, Nanjing, 210096, China (e-mail: changhengtai@pmlabs.com.cn).

Ji Bian is with School of Information Science and Engineering, Shandong Normal University, Jinan, 250399, China (e-mail: jibian@sdnu.edu.cn).

methods do not need information on the cluster number. However, it tends to focus only on the distribution of MPCs in the delay domain, lacking angular information. In [25], [26], a clustering approach based on image processing was proposed, such as the Hough-transform algorithm and power-angle-spectrum-based clustering and tracking (PASCT) algorithm. In addition to the aforementioned traditional algorithms, clustering algorithms based on the Gaussian mixture model (GMM) have been increasingly utilized for clustering MPCs. GMM-based algorithms have shown more excellent suitability for channel clustering due to their ability to incorporate more statistical features. In [27], [28], GMM was used for clustering in O2I and HSR scenarios. However, there is limited literature available that employs this more recent approach.

Furthermore, clusters may exhibit complex evolution behaviors, such as birth/death, movement, splitting, or merging. In [29], an automatic tracking algorithm has been proposed to track the MPCs in vehicle-to-vehicle (V2V) channels. In [30], a distance-based tracking algorithm was proposed to analyze the cluster lifetime, and the density-based spatial clustering of applications with noise (DBSCAN) algorithm was used for channel clustering. Besides, multipath component distance (MCD) based tracking methods were applied in [31], [32]. These studies have shown that this MCD-based tracking method can obtain better robustness and accuracy when dealing with time-varying channel data.

For UAV-to-ground communications, the investigation of channel clustering is still in its early stages. The clustering and tracking of MPCs for UAV-to-ground communication channels was presented in [33]. The potential MPCs were estimated using the space-alternating generalized expectation maximization (SAGE) algorithm based on measurements at 6.5 GHz with 500 MHz bandwidth. Clustering methods and modeling approaches are conventional. Subsequently, the extracted MPCs were clustered using the traditional KPM algorithm. Besides, the cluster-based tracking (CBT) method was used to quantify the survival length of the clusters. However, it only considers information about the delay domain and lack of angle information. Besides, it does not involve millimeter-wave (mmWave) bands. In [34], the high-resolution-parameter-estimation (HRPE) principle was used to obtain MPCs, and the clustering method based on the MCD threshold was applied to briefly analyze the channel characteristics. It also only focused on the sub-6 GHz and used traditional clustering methods.

In general, current models primarily concentrate on conventional terrestrial channels, such as indoor, V2V, and HSR channels. However, the UAV channel has significant differences from conventional terrestrial channels, for example, large elevation angles, arbitrary trajectories, etc. It leads to significant differences in inter-cluster and intra-cluster characteristics such as the number, statistical distribution, and lifetime of clusters. Therefore, it is urgent to conduct appropriate clustering algorithm research and analyze the relevant characteristics of UAV-to-ground communication channels.

Motivated by the above background and gaps of current research, a novel ML-based clustering and modeling method for 6G UAV-to-ground communication channels is proposed. To the best of our knowledge, this is the first study that concentrates on the clustering of the UAV-to-ground communication channels at both sub-6 GHz and mmWave bands. The main contributions and novelties of this paper can be summarized as follows.

1) Abundant time-varying UAV-to-ground communication channel datasets at both sub-6 GHz and mmWave are acquired by the ray-tracing (RT) modeling method, and the VB-GMM algorithm is used to conduct channel clustering. The algorithm automatically determines the optimal cluster number and considers the delay and space domains.
2) The inter- and intra-cluster characteristics of UAV-to-ground communication channels are comprehensively analyzed, including the number of clusters, cluster delay, cluster power, intra-cluster angular and delay offset, etc. The MCD-based tracking method is applied to capture the evolving behavior of time-varying clusters, and the lifetime of clusters is also analyzed.
3) A cluster-based UAV-to-ground communication channel model is proposed, and a four-state Markov chain is introduced to portray the evolution of clusters. Some typical channel characteristics are studied based on the proposed channel model, such as delay power spectrum densities (PSDs) and root-mean-square (RMS) delay spread (DS). Besides, the results of the channel measurements are consistent with the simulations, verifying the effectiveness of the model.

The remainder of this paper is organized as follows. Section II presents the channel clustering and tracking algorithm. In Section III, the simulation environment and clustering results are illustrated. Section IV presents the intra-cluster, inter-cluster, and time-varying evolution characteristics. Then, a cluster-based channel model is proposed and implemented in Section V. Finally, conclusions are drawn in Section VI.

Notation: Through the whole manuscript, $(\cdot)^T$ represents the transpose, $\mathrm{tr}(\cdot)$ denotes the trace operation, $(\cdot)^{-1}$ represents the inversion, $\|\cdot\|$ denotes the Euclidean norm, $\det(\cdot)$ denotes the determinant, and $|\cdot|$ represents the absolute value.

II. Channel Clustering and Time-Varying Clusters Tracking Algorithm

A. Description of Clustering Problem

Fig. 1 illustrates the clustering and evolution phenomenon of MPCs in UAV-to-ground communications. Different types of scattering clusters often affect radio propagation. For instance, the ground station (GS) usually detects signals from various propagation paths, such as tall buildings, bungalows, ground, trees, vegetation, etc. In this scenario, the UAV acts as the transmitter (Tx), while the GS consists of the ground user or base station (BS) and serves as the receiver (Rx). Due to its high flight altitude, UAV can transmit signals directly from the Tx to Rx, which is also an advantage of UAV communication [2]. Therefore, the communication path always contains the line-of-sight (LoS) component, which is considered as LoS situations.

In this paper, the 6-dimensional (6-D) input data, i.e., the delay, elevation angle of arrival (EAOA), elevation angle of

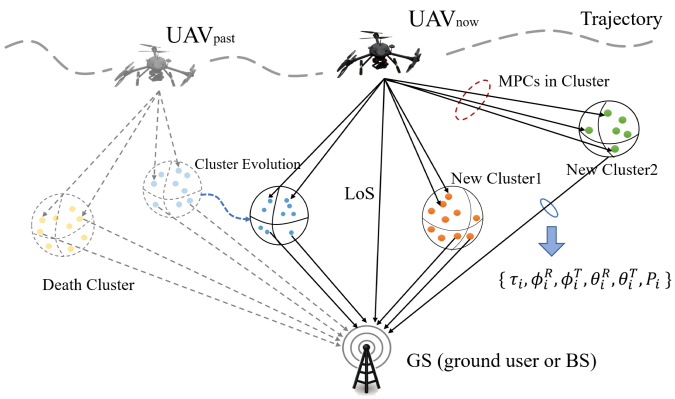


Fig. 1. The architecture of clustering and evolution phenomenon for UAV-to-ground communication channels.

departure (EAOD), azimuth angle of arrival (AAOA), azimuth angle of departure (AAOD), and power are used to implement channel clustering. The 6-D channel data of the i-th MPC correspond to τ_i, θ_i^R, θ_i^T, ϕ_i^R, ϕ_i^T, and P_i, respectively. Our target is to obtain the cluster labels for each MPC based on the similarity between these 6-D parameters of the MPCs. Furthermore, MPCs can be classified into separate clusters, and the optimal number of categories is obtained. Additionally, note that the following clustering method has excluded the LoS path and only contains quantization for the non-line-of-sight (NLoS) paths. Besides, due to the significant power difference between the LoS path and other paths, we first extract the LoS path from the multipath components and then cluster the NLoS paths. Moreover, it is more in accordance with the idea of standardized channel modeling [13]–[15].

B. VB-GMM Clustering Algorithm

With the increasing bandwidth of communication systems, the diversity of scenarios, and the high mobility of terminals, channel multipath presents big data characteristics and also has a statistical distribution pattern. Traditional distance-based clustering methods, such as K-means and FCM, can be ineffective in mining the statistical distribution pattern. In this paper, a GMM-based clustering method is chosen. Due to differences in the size of the 6-D channel raw data, the range has been rescaled to [-1, 1] before clustering. Therefore, we can get the normalized channel data matrix $\boldsymbol{X}$, which contains some sub-data $\boldsymbol{x}$. The channel data is then inverse normalized after the clustering has been completed.

The GMM assumes that each sub-data $\boldsymbol{x}$ is generated based on the weight vector π_i and the Gaussian probability distribution function (PDF) of i-th component. The obtained mixture model is given by [35]

$$p(\boldsymbol{x}) = \sum_{i=1}^{K} \pi_i \cdot N(\boldsymbol{x}; \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \tag{1}$$

$$N(\boldsymbol{x}; \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) = \frac{\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu}_i)^T \boldsymbol{\Sigma}_i^{-1}(\boldsymbol{x}-\boldsymbol{\mu}_i)\right)}{\sqrt{(2\pi)^d \det(\boldsymbol{\Sigma}_i)}} \tag{2}$$

where $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}_i$ are the mean and covariance matrix of the i-th Gaussian component, d is the dimension, and K is the total number of components.

Therefore, the posterior probability can be calculated as

$$p(i \mid \boldsymbol{x}) = \frac{\pi_i \cdot N(\boldsymbol{x} \mid \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_\iota)}{\sum_{\iota=1}^{K} \pi_\iota \cdot N(\boldsymbol{x} \mid \boldsymbol{\mu}_\iota, \boldsymbol{\Sigma}_\iota)}. \tag{3}$$

In order to cluster all MPCs, it is necessary to determine the posterior probability, also known as responsibility. The responsibility $p(i|\boldsymbol{x})$ represents the probability that each MPC is generated by the i-th Gaussian distribution. Each MPC is assigned to the Gaussian distribution with the highest posterior probability.

Finding responsibility is equivalent to solving a hidden-variable problem. The parameters required for this solution are $\{\pi_i, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}$. It is usually solved by iterative optimization using the Expectation Maximization (EM) algorithm. However, the EM algorithm requires cross-validation to determine the optimal cluster, and the covariance matrix is probably singular. In this paper, the VB method is used to solve the problem, which automatically determines the optimal number of clusters without the assistance of cross-validation and has good convergence. For convenience, the covariance matrix $\boldsymbol{\Sigma}_i$ in (2) has been rewritten as the precision matrix $\boldsymbol{T}_i$.

The Bayesian GMM can be obtained by applying prior probability distributions to the parameters $\Theta = \{\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{T}\}$. The Dirichlet prior for $\boldsymbol{\pi}$ and the Gauss-Wishart prior for $(\boldsymbol{\mu}, \boldsymbol{T})$ are introduced. The Dirichlet prior is used for $\boldsymbol{\pi}$ with the parameters α_i as [27]

$$\mathrm{Dir}(\boldsymbol{\pi} \mid \{\alpha_i\}) = \frac{\Gamma\left(\sum_{i=1}^{K} \alpha_i\right)}{\prod_{i=1}^{K} \Gamma(\alpha_i)} \cdot \prod_{i=1}^{K} \pi_i^{\alpha_i - 1} \tag{4}$$

where $\Gamma(\cdot)$ is the Gamma function. The parameter α_i can be interpreted as the number of valid prior observations associated with each component of the mixture. The Gaussian-Wishart prior for $(\boldsymbol{\mu}, \boldsymbol{T})$ is consists of the Gaussian distribution $N(\cdot)$ and Wishart distribution $\mathcal{W}(\cdot)$ as

$$p(\boldsymbol{\mu}, \boldsymbol{T}) = \prod_{i=1}^{K} N(\boldsymbol{\mu}_i; \boldsymbol{\mu}_0, \beta_0 \boldsymbol{T}_i) \cdot \mathcal{W}(\boldsymbol{T}_i \mid v, \boldsymbol{V}) \tag{5}$$

where $\boldsymbol{V}$ and v represent the scale matrix and the degrees of freedom, respectively. The Wishart distribution is a generalization of the Gamma distribution in multiple dimensions, and the Wishart prior is applied to account for the possible correlations.

From the prior probability distributions to the parameters Θ, it is evident that this Bayesian GMM only relies on hidden random variables, namely $h = (\boldsymbol{Z}, \boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{T})$. Using the variational methodology, an approximation $q(h)$ is computed for the Bayesian GMM, which is expressed as a product of the form as [27]

$$q(h) = q(\boldsymbol{Z}) \cdot q(\boldsymbol{\pi}) \cdot q(\boldsymbol{\mu}, \boldsymbol{T}). \tag{6}$$

Algorithm 1 Clustering Algorithm for UAV-to-Ground Communication Channels.

Input: The normalized channel matrix $\boldsymbol{X}$ and the total number of sampling points S_p.
Output: The labels of MPCs and the number of cluster.

```
for snapshot = 1: S_p do
    Initialize the prior parameters, and the maximum iteration number M_iter;
    while iter < M_iter do
        calculate the responsibility of expectation (Eq. 7);
        apply current responsibility values to update model parameter (Eq. 8–11);
        calculate the difference of variational lower bound;
        if the convergence condition is satisfied then
            label = max_i{p(i|x)};
            calculate the total number of labels;
            break;
        else
            iter ++;
        end if
    end while
end for
```

The details on the updating of the parameters and calculation process are referred in [35], and the result is as follow [28]:

$$q(Z) = \prod_{n=1}^{N}\prod_{i=1}^{K} r_{in}^{z_{in}} \tag{7}$$

$$q(\boldsymbol{\pi}) = \mathrm{Dir}\left(\boldsymbol{\pi} \mid \{\alpha_i\}\right) \tag{8}$$

$$q(\boldsymbol{\mu}, \boldsymbol{T}) = \prod_{i=1}^{K} q\left(\boldsymbol{T}_i\right) \cdot q\left(\boldsymbol{\mu}_i \mid \boldsymbol{T}_i\right) \tag{9}$$

$$q(\boldsymbol{T}) = \prod_{i=1}^{K} \mathcal{W}\left(\boldsymbol{T}_i; \eta_i, \boldsymbol{U}_i\right) \tag{10}$$

$$q\left(\boldsymbol{\mu}_i \mid \boldsymbol{T}_i\right) = \prod_{i=1}^{K} N\left(\boldsymbol{\mu}_i; m_i, \beta_i \boldsymbol{T}_i\right) \tag{11}$$

where η_i and $\boldsymbol{U}_i$ are parameters of the Wishart distribution representing the degrees of freedom and scale matrix, respectively. Then, the best variational lower bound $L(\mathrm{iter})$ can be obtained by performing several iterations. Besides, the best approximation $q(h)$ to the true posterior $p(h|\boldsymbol{x})$ can be obtained by the iterative method [27]. Therefore, the label of each MPC can be obtained. In summary, the detailed relevant pseudo-code is shown in Algorithm 1. The input dataset $\boldsymbol{X}$ is the normalized 6-D channel data in all snapshots. Each *for* loop output is the clustering result under each snapshot, each *while* loop is to find whether the iteration requirement is satisfied, and finally channel clustering results for all snapshots is obtained.

This paper uses the synthetic dataset to better illustrate the performance of the VB-GMM method [28]. In the experiment, a total of 8000 data points in three-dimensional (3-D) space are randomly generated from the four clusters, and each cluster has 2000 sample points. The mean matrices of the generated matrix are $(-1,-2,-2)^T$, $(3,-2,2)^T$, $(-3,1,-2)^T$, and $(3,1,2)^T$, respectively. For the covariance matrix, it is required that the matrix is semi-positive definite [35]. Thus, the generated covariance matrices is $\begin{pmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \end{pmatrix}$, $\begin{pmatrix} 2 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 4 \end{pmatrix}$, $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, and $\begin{pmatrix} 2 & 0 & -0.5 \\ 0 & 1 & 1 \\ -0.5 & 1 & 2 \end{pmatrix}$.

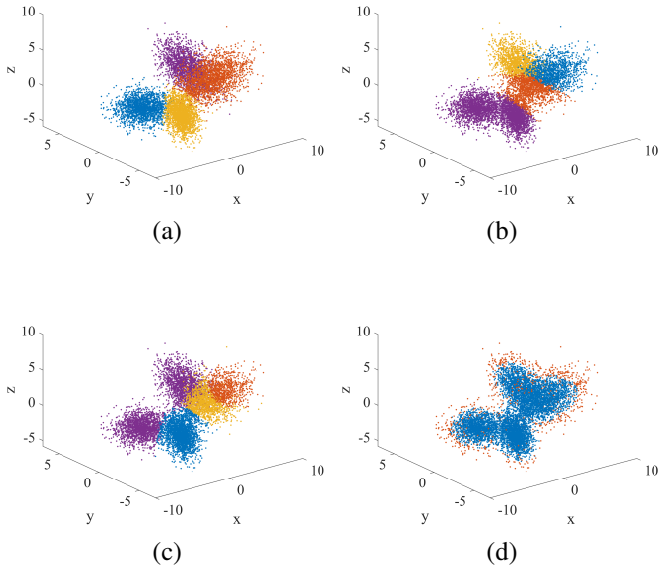


Fig. 2. Synthetic dataset clustering results of (a) VB-GMM, (b) K-means, (c) Spectral, and (d) DBSCAN clustering algorithms (the same colour represents one cluster).

Fig. 2 shows the clustering results of typical clustering methods, i.e., VB-GMM, K-means, Spectral, and DBSCAN algorithms. The same colors symbolize the same categories. Due to random phenomena such as initialization, the random seed of 4 is firstly set to facilitate the display in this paper. It is difficult to separate the non-circular distribution for the determination based on distance, such as the K-means algorithm in Fig. 2(b). Besides, the K-means algorithm here sets the number of clusters in advance to 4. The points that are obviously different from each other but closer are often regarded as one class. The identified clusters by the Spectral and DBSCAN clustering methods are also confused, as shown in Fig. 2(c) and (d). These approaches are less effective for the data with Gaussian-like distribution properties, although they do not require the number of clusters to be known in advance. In addition, the K-means, Spectral, and DBSCAN clustering algorithms are susceptible to falling into local optimal solutions. The VB-GMM-based clustering algorithm works better because it considers the overall information and can be applied to any distribution theoretically.

Also, the effectiveness evaluation indexes of the relevant clustering algorithms are calculated. The evaluation indexes can be divided into internal and external effectiveness in-

TABLE I
THE CLUSTERING ALGORITHM PERFORMANCE INDICES FOR SYNTHETIC DATASETS

Methods	Internal effectiveness indexes			External effectiveness indexes			
	DBI	CH	SC	NMI	ARI	JI	FMI
VB-GMM	1.17	**7659**	0.51	**0.83**	**0.87**	**0.90**	**0.90**
K-means	**0.99**	6523	**0.52**	0.60	0.49	0.70	0.63
Spectral	1.42	2052	0.11	0.58	0.50	0.74	0.65
DBSCAN	5.42	125	0.34	0.02	0.004	0.82	0.45

dexes. The internal effectiveness indexes mainly evaluate the tightness, connectivity, and overlap of the clustering result, including Silhouette coefficient (SC), Davies-Bouldin index (DBI), and Calinski-Harabasz index (CHI). External effectiveness indexes are obtained by comparing the match between the clustering results and external guidelines, including normalized mutual information (NMI), adjusted rand index (ARI), Jaccard index (JI), and Fowlkes-Mallows index (FMI). The formulas for the related indicators can be found in [35], [36]. Note that smaller values for the results of the DBI imply better clustering performance, while the opposite is the CH, SC, NMI, ARI, JI, and FMI. Table I shows the results of the evaluation indicators for the different clustering methods. The bolded ones in each column represent the best performance. Overall, the VB-GMM method presents better performance for most of the indexes. The best results are presented for the NMI, ARI, JI, FMI, and CH. The DBI and SC performance is ranked second, not much different from the first. With the help of the above indicators, it can be found that VB-GMM performs better, overcomes the disadvantages of other methods, and can obtain the cluster number automatically.

C. Time-Varying Cluster Tracking

For UAV-to-ground communications, clusters often undergo birth-death (B-D) phenomena. Some indicators of the mobile communication systems can be changed during flight. Therefore, the evolution of the clusters needs to be depicted. In previous methods, the Euclidean distance (ED) was often used to measure the distance between components. However, for the channel dataset, the data units of each parameter of the MPCs are inconsistent, and the direct use of the ED can not provide objective and accurate tracking results. Moreover, due to the fast cluster change in UAV-to-ground communications, considering only the two closest sampling points often results in interrupted trajectories, which is unsuitable for further analysis. In this paper, a tracking method based on MCD is adopted. Here, the MCD can be calculated as [23]

$$\mathrm{MCD}_{ij} = \sqrt{(\mathrm{MCD}_{ij}^{\tau})^2 + (\mathrm{MCD}_{ij}^{\mathrm{AoD}})^2 + (\mathrm{MCD}_{ij}^{\mathrm{AoA}})^2}. \tag{12}$$

The MCD of delay is given by

$$\mathrm{MCD}_{ij}^{\tau} = \zeta \cdot \frac{|\tau_i - \tau_j|}{\Delta\tau_{\max}} \cdot \frac{\tau_{\mathrm{std}}}{\Delta\tau_{\max}} \tag{13}$$

where τ_{std} is the standard deviation (std) of the MPCs delay, $\Delta\tau_{\max}$ is the maximum value of the MPCs delay difference, and ζ is the weight of the delay distance.

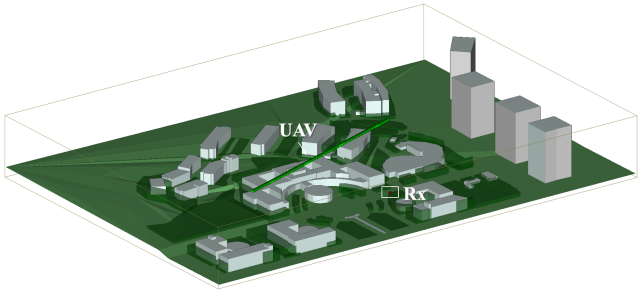


Fig. 3. The RT-based reconstructed environments.

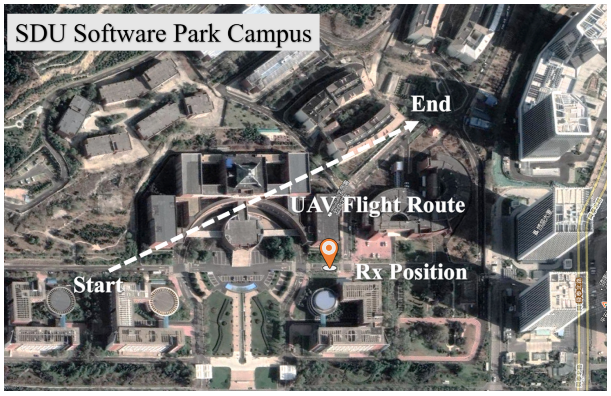


Fig. 4. The satellite view with flight detail sketch for UAV communication scenarios.

The MCD of angular is shown as

$$\mathrm{MCD}_{ij}^{\mathrm{AoA/AoD}} = \frac{1}{2} \cdot \left\| \begin{pmatrix} \sin(\theta_i^{\mathrm{R/T}})\cos(\phi_i^{\mathrm{R/T}}) \\ \sin(\theta_i^{\mathrm{R/T}})\sin(\phi_i^{\mathrm{R/T}}) \\ \cos(\theta_i^{\mathrm{R/T}}) \end{pmatrix} - \begin{pmatrix} \sin(\theta_j^{\mathrm{R/T}})\cos(\phi_j^{\mathrm{R/T}}) \\ \sin(\theta_j^{\mathrm{R/T}})\sin(\phi_j^{\mathrm{R/T}}) \\ \cos(\theta_j^{\mathrm{R/T}}) \end{pmatrix} \right\|. \tag{14}$$

Moreover, the threshold in delay and the angular domain is set to 0.2 μs and 0.5 rad, respectively. Due to the delay and angular domain variations of clusters do not change significantly considering the actual situation [33]. If the threshold is beyond, the cluster is regarded as death. Similarly, if the past cluster center can not be captured, the cluster is considered newly generated.

III. CHANNEL DATASETS ACQUISITION AND CLUSTERING RESULT

A. RT-Based Channel Model Datasets

RT is a deterministic channel modeling method based on geometrical optics (GO) and uniform theory of diffraction (UTD). It considers the interaction between rays and scatterers, including reflection, scattering, and diffraction phenomena. By tracing all possible ray paths, the MPCs between the Rx and Tx can be determined. This enables an accurate characterization of the overall propagation process. Besides, RT can generate a large amount of data with high accuracy and can match the actual channel data well [37].

The Software Park Campus of Shandong University (SDU) in Jinan, China, was chosen for the UAV communication

TABLE II
THE SIMULATION DETAILS AND PARAMETERS OF RAY-TRACING MODEL

Material Parameters	Name	Type	Permittivity	Conductivity (S/m)	Thickness (m)
	concrete	one-layer dielectric	7.000	0.0150	0.300
	glass	one-layer dielectric	6.270	0.2287 / 0.1915	0.003
	wet earth	dielectric half-space	25.000	0.0200	—
	vegetation	tree	Leaf radius (m)	Leaf thickness (m)	Leaf density
			0.050	0.0005	250
		grass	Blade radius (m)	Blade length (m)	Blade disity /area
			0.002	0.460	900
Physical Parameters	Tx altitude	Rx altitude	Flight speed	Flight distance	Sampling interval
	50 m	2 m	10 m/s	300 m	1 m
Simulation Parameter	Frequency	Bandwidth	Transmit power	Antenna pattern	Antenna max gain
	28 GHz / 3.8 GHz	1 GHz / 160 MHz	10 dBm	isotropic	0 dBi
	Polarization	Waveform	Threshold	Ray-spacing	Reflection/Diffraction/Transmission
	vertical	sinusoid	−160 dBm	0.25	6/0/1

scenarios. The channel propagation scenarios were constructed using Wireless InSite software [38]. Fig. 3 shows the reconstructed simulation scenarios, and Fig. 4 shows the satellite view with flight detail sketch. The scenario includes various buildings, vegetation, glass, etc, and the topography terrain was also taken into account. The category of vegetation is set to Biophysical. The height of trees is set to 5–12 m according to the actual scenario, the height of grass is set to 0.2 m, and the height of buildings is set to approximately 15–70 m. The overall layout of the buildings is complex. The material is set to glass in some positions according to the actual situation. The total area is approximately 650×470×80 m^3. The selection of material types and electromagnetic parameters are referred to the International Telecommunication Union (ITU) database [38], [39]. Moreover, we also introduced the Lambertian diffuse reflection model to suit the actual radio propagation. The UAV carries the Tx while the ground serves as the Rx, and the UAV flight trajectory is also marked in Fig. 4, crossing over many buildings. The simulation details and parameter selection settings are shown in Table II.

B. Clustering Results

The proposed VB-GMM clustering algorithm can cluster MPCs in both space and time domains, also considering power influence. Fig. 5 shows the clustering results of MPCs in sub-6 GHz and mmWave bands for one snapshot. The delay, AAOA, and AAOD are selected for visualization to facilitate the analysis, where the same clusters have the same color. It is observed that both sub-6 GHz and mmWave bands have effective clustering. As the mean and variance information is considered, the clustering results correspond well to the multipath propagation characteristics.

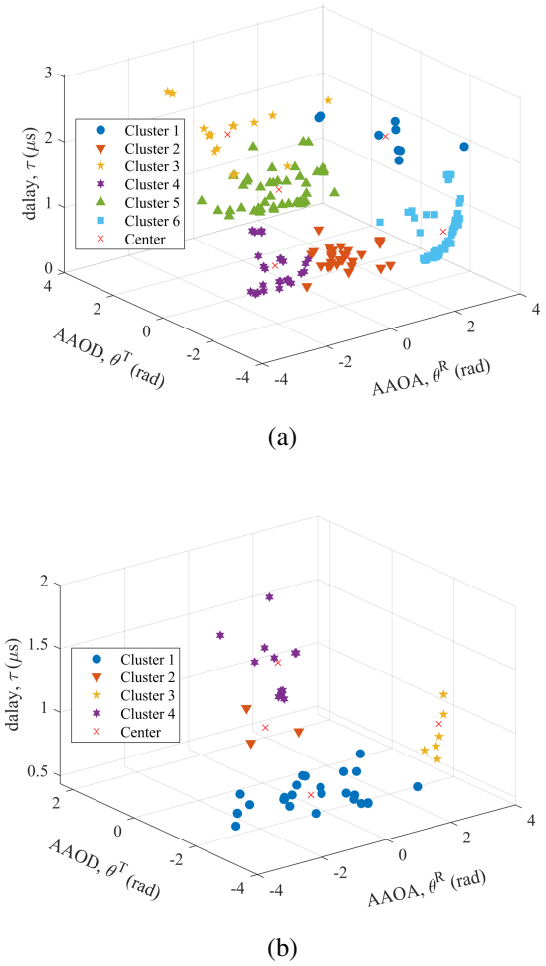


Fig. 5. The clustering results in the visual aspect in (a) sub-6 GHz band and (b) mmWave band.

IV. CHANNEL CLUSTERING CHARACTERIZATION ANALYSIS FOR UAV-TO-GROUND COMMUNICATIONS

A. Intra-Cluster Characterization

The results of channel clustering provide the cluster label of each MPC as well as the total number of clusters. Further computation allows us to obtain the intra-cluster characterization. Table III summarizes the statistical modeling results of the intra-cluster characteristics. The details are as follows.

1) The Number of Intra-Cluster MPCs: Fig. 6 shows the cumulative distribution functions (CDFs) of the number of intra-cluster MPCs. The normal distribution is a suitable model for the number of intra-cluster MPCs. In sub-6 GHz band, the mean value of the number of intra-cluster MPCs is 29.73 with the std value of 28.44, while in mmWave band, the mean value is 10.36 with the std value of 8.74. The number of intra-cluster MPCs in sub-6 GHz is larger than that in mmWave. This is

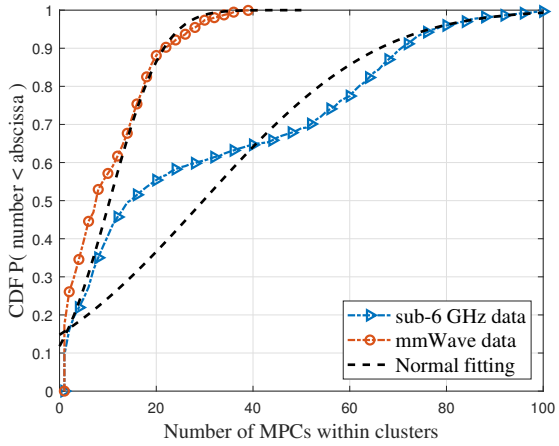


Fig. 6. The CDFs of the number of MPCs within cluster at different bands.

because the path loss of MPCs in mmWave is more significant under the same received threshold, making it difficult to detect more intra-cluster MPCs at the Rx end.

2) Intra-Cluster Power Decay Factor: It is well known that the intra-cluster MPCs exhibit a linear sloped decaying relationship for the power in dB with delay in μs, and different numbers of MPCs tend to have different decreasing trends [33]. Our desire is to build the relationship between the intra-cluster power decay factor k and the MPCs number within cluster C. Their relationship can usually be modeled by exponential stochastic processes, whereas first-order exponential (exp1) models are not sufficiently refined. Therefore, we use the second-order exponential (exp2) model to picture this relationship. It can be found that the exp2 model is better thought out in both of the bands than exp1 model for relevant data indicators, i.e., root mean square error (RMSE) and R-square, and exp2 model can portray its relationship well. Table III represents the relationship between intra-cluster power decay factor and MPCs number within cluster for different bands, and gives the values for exp2 model in details.

3) Cluster Rician K-factor: To better characterize the power distribution of the intra-cluster MPCs, the cluster Rician K-factor is adopted for modeling, and can be calculated as

$$\Lambda(\mathrm{dB}) = 10 \cdot \log_{10}\left(\frac{\max(P_c)}{\sum_{c=1}^{C} P_c - \max(P_c)}\right) \tag{15}$$

where C is the MPCs number within cluster and Λ is cluster Rician K-factor in dB. Table III demonstrates that the normal distribution effectively models the cluster Rician K-factor. The mean values are 5.11 dB in sub-6 GHz and mmWave bands, while the std values are 9.07 dB and 6.49 dB, respectively.

4) Intra-Cluster Delay Distribution: The cluster RMS DS describes the time-dispersion characteristics of the MPCs in UAV-to-ground communication channels. The intra-cluster RMS DS is expressed as ς, and its calculation method is

$$\varsigma = \sqrt{\frac{\sum_{c=1}^{C} P(\tau_c) \cdot \tau_c^2}{\sum_{c=1}^{C} P(\tau_c)} - \left(\frac{\sum_{c=1}^{C} P(\tau_c) \cdot \tau_c}{\sum_{c=1}^{C} P(\tau_c)}\right)^2}. \tag{16}$$

The intra-cluster RMS DS follows the log-normal distribution, i.e., $\ln(\varsigma) \sim \mathcal{N}(\mu_{DS}, \sigma_{DS})$. The unit of ς is μs.

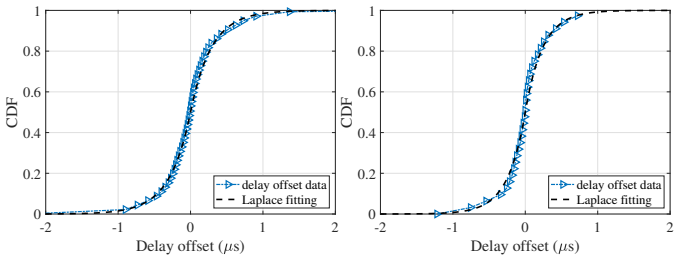


Fig. 7. The CDFs of delay offset in (a) sub-6 GHz band and (b) mmWave band.

The mean and std parameters for the log-normal distribution of intra-cluster RMS DS for mmWave are -2.56 and 0.65, respectively. Meanwhile, the mean and std parameters for sub-6 GHz are -1.98 and 0.46, respectively. These findings indicate that delay extension within the cluster is insignificant. The subpaths within the clusters have comparable delay, which also demonstrates the efficacy of the clustering outcomes. Additionally, the intra-cluster is more spread out in sub-6 GHz compared to mmWave band.

Also, the delay offset is an essential factor in generating the delay of subpaths in a cluster. It is defined as the difference between the delay of the subpaths in the cluster and the average delay of the cluster, i.e., $\tau_{\mathrm{off}} = \tau_c - \mathrm{mean}(\tau_c)$. Fig. 7 shows the CDFs of delay offset and the corresponding fitting results for sub-6 GHz and mmWave. It is found that the Laplace distribution with zero mean can model the delay offset well [40]. The scale parameter χ_τ determines the width of the delay offset distribution, which is calculated as the std value divided by $\sqrt{2}$. Larger scale parameters imply wider distributions and more disperse. In this paper, the scale parameters χ_τ are 0.29 and 0.24 for sub-6 GHz and mmWave bands.

5) Intra-Cluster Angular Distribution: Similar to the intra-cluster MPCs delay characteristics, we use the RMS angular spreads (ASs) and angular offsets to characterize the angular distribution of the MPCs within the cluster. The log-normal distribution is used to fit intra-cluster RMS ASs, and the specific parameters are in Table III. Fig. 8 shows the CDFs of cluster RMS ASs and the corresponding fitting results at different bands. It can be found that cluster ASs are similar in both bands. In terms of mean value and 80% share of RMS ASs, both are consistent with the azimuth angular of arrival spread (AAS) $>$ the azimuth angular of departure spread (ADS) $>$ the elevation angular of arrival spread (EAS) $>$ the elevation angular of departure spread (EDS). This is because the azimuth in the cluster comes from the various distributions of buildings around the cluster, which results in a decentralized azimuth distribution. Similarly, the elevation angular is mainly influenced by the height of the UAV and the ground Rx, resulting in a concentrated angular. In addition, the arrival ASs are larger than the departure ASs because the scatterers are mainly distributed at the ground end. Also, the angular offset is introduced to generate the angle of the subpaths in the cluster. Fig. 9 and 10 show the CDFs of angular offsets and the corresponding fitting results. The detailed parameters

TABLE III
INTRA-CLUSTER CHANNEL CHARACTERIZATION PARAMETERS

Channel parameter		Statistical distribution	Value	
			sub-6 GHz @ 3.8 GHz	mmWave @ 28 GHz
Number of intra-cluster MPCs		Normal distribution [mean μ_N, std σ_N]	μ_N = 29.73, σ_N = 28.44	μ_N = 10.36, σ_N = 8.74
Power decay factor		Second-order exponential distribution $[a_1 \exp(b_1 x) + a_2 \exp(b_2 x)]$	$a_1 = -32.15$, $b_1 = -0.45$, $a_2 = -3.62$, $b_2 = -0.03$	$a_1 = -0.72$, $b_1 = 0.01$, $a_2 = -9.05$, $b_2 = -0.12$
Cluster Rician K-factor (dB)		Normal distribution [mean μ_{KF}, std σ_{KF}]	μ_{KF} = 5.11, σ_{KF} = 9.07	μ_{KF} = 5.11, σ_{KF} = 6.49
Cluster delay distribution	RMS DS (μs)	Log-normal distribution [mean μ_{DS}, std σ_{DS}]	$\mu_{DS} = -1.98$, σ_{DS} = 0.46	$\mu_{DS} = -2.56$, σ_{DS} = 0.65
	delay offset (μs)	Zero-mean Laplace distribution [scale χ_τ]	χ_τ = 0.29	χ_τ = 0.24
Cluster angular distribution	RMS AAS (rad)	Log-normal distribution [mean μ_{AAS}, std σ_{AAS}]	$\mu_{AAS} = -1.01$, σ_{AAS} = 0.41	$\mu_{AAS} = -1.01$, σ_{AAS} = 0.58
	RMS EAS (rad)	Log-normal distribution [mean μ_{EAS}, std σ_{EAS}]	$\mu_{EAS} = -2.35$, σ_{EAS} = 0.48	$\mu_{EAS} = -2.54$, σ_{EAS} = 0.60
	RMS ADS (rad)	Log-normal distribution [mean μ_{ADS}, std σ_{ADS}]	$\mu_{ADS} = -1.67$, σ_{ADS} = 0.46	$\mu_{ADS} = -1.91$, σ_{ADS} = 0.49
	RMS EDS (rad)	Log-normal distribution [mean μ_{EDS}, std σ_{EDS}]	$\mu_{EDS} = -2.91$, σ_{EDS} = 0.43	$\mu_{EDS} = -2.96$, σ_{EDS} = 0.54
	AAOA offset rad)	Zero-mean Laplace distribution [scale b_{AAOA}]	b_{AAOA} = 0.73	b_{AAOA} = 0.76
	AAOD offset (rad)	Zero-mean Laplace distribution [scale b_{AAOD}]	b_{AAOD} = 0.12	b_{AAOD} = 0.13
	EAOA offset (rad)	Zero-mean Laplace distribution [scale b_{EAOA}]	b_{EAOA} = 0.38	b_{EAOA} = 0.34
	EAOD offset (rad)	Zero-mean Laplace distribution [scale b_{EAOD}]	b_{EAOD} = 0.10	b_{EAOD} = 0.10

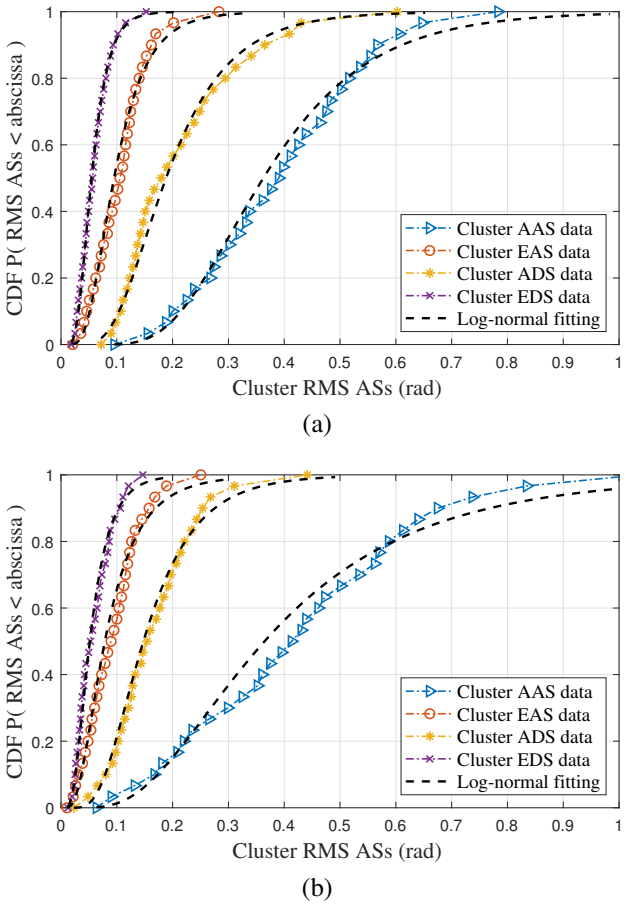


Fig. 8. The CDFs of cluster RMS ASs at different bands in (a) sub-6 GHz band and (b) mmWave band.

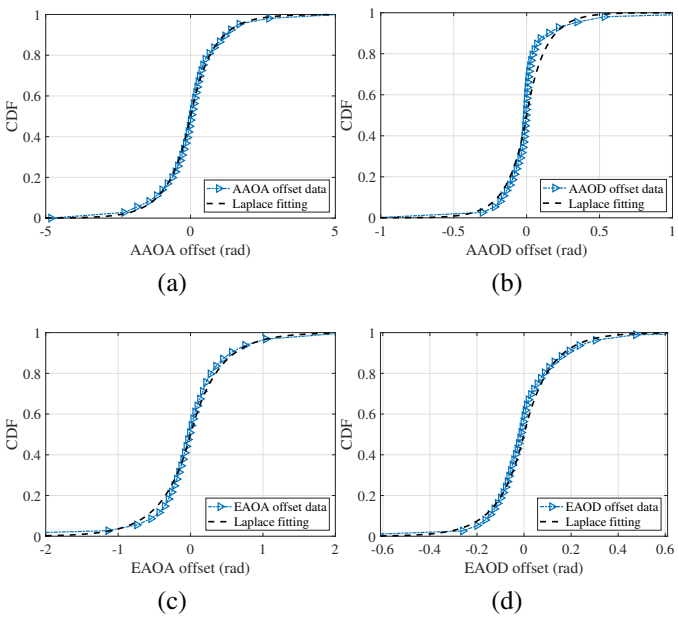


Fig. 9. The CDFs of (a) AAOA offset, (b) AAOD offset, (c) EAOA offset, and (d) EAOD offset in sub-6 GHz band.

can also be found in Table III.

B. Inter-Cluster Characterization

The inter-cluster parameter also reflects the characteristics of the channel between clusters. The details are as follow.

1) The Number of Clusters: The number of clusters also reflects the abundance of clusters in UAV-to-ground communication channels, which is related to the communication

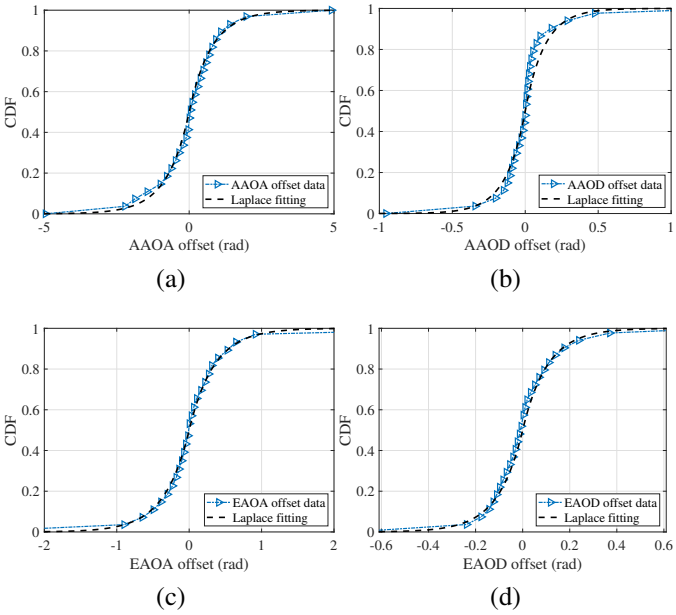


Fig. 10. The CDFs of (a) AAOA offset, (b) AAOD offset, (c) EAOA offset, and (d) EAOD offset in mmWave band.

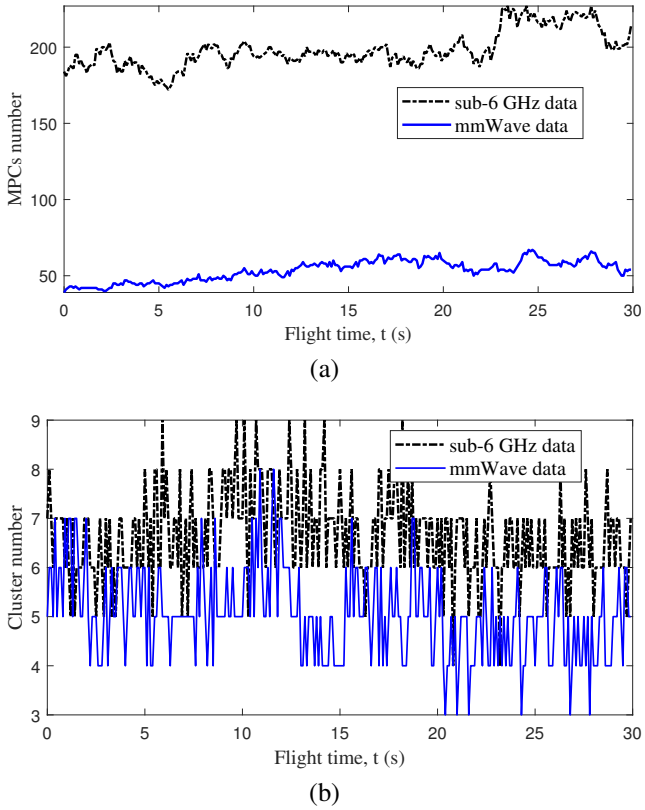


Fig. 11. The time-varying results of (a) MPCs number and (b) cluster number.

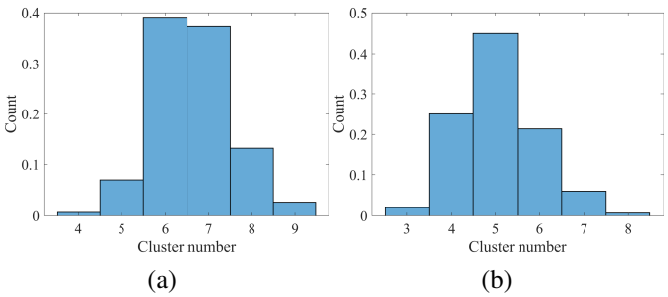


Fig. 12. The count histogram of the cluster number in (a) sub-6 GHz band and (b) mmWave band.

environment and the frequency band. Fig. 11 shows the number of clusters and MPCs both of the bands based on the clustering results. The number of MPCs and clusters have a similar relationship. The maximum number of clusters at sub-6 GHz band can be up to 9, and the maximum number at mmWave band can be up to 8. Besides, the number of MPCs at mmWave band is relatively small. The number of clusters is counted as shown in Fig. 12. The maximum probability of the number of clusters at sub-6 GHz is 6, accounting for 39%. The maximum probability of the number of clusters at mmWave is 5, accounting for 45%.

2) Inter-Cluster Power and Delay Characteristics: Based on the different clusters obtained, the power and delay are divided. Table IV gives the relative delay and power for different clusters. In order to eliminate the effect of the LoS path at different locations, the power and delay are derived by comparing them with the delay and power of the LoS paths [33]. The power is significantly lower in mmWave band in simulation. Therefore, fewer MPCs are perceived at the Rx end, and it can be modeled with fewer clusters. Besides, compared to the traditional terrestrial channel model, the UAV-to-ground communication channels have more extensive delay range, which indicates that the effect of tall buildings in the distance on propagation can be perceived. Moreover, the delay and power of cluster index of 3–5 are relatively similar, which demonstrates that these may be coming from buildings with different orientations but with similar relative distances.

3) Inter-Cluster Angular Distribution: The space distribution of cluster locations is observed. The cluster appearance angles are similar for sub-6 GHz and mmWave bands. The AAOA is approximately uniformly distributed in the range $[-180°, 180°)$. The EAOA is mainly distributed in [85°, 100°], with sporadic distribution in other places greater than 100°. The AAOD is distributed in the range of $[-180°, 50°]$. The main reason is that the high buildings are distributed in this range, so the first bounce is concentrated in this angle range. The EAOD is mainly distributed in [90°, 130°] and shows a decreasing trend. The key reasons for this are the three relative positions of the scatterers, the Rx and UAV, and the distribution of the scatterers in the environment. Overall, it can be assumed that the distribution of scatterers in the horizontal plane is approximately uniform. This makes AAOA and AAOD also more uniformly distributed over the whole angle. In the vertical dimension, scatterers are mainly distributed at the near-ground end, e.g., vegetation, bungalows, and trees. Thus, EAOA and EAOD are distributed over a range of angles and concentrated near [85°, 130°].

TABLE IV
INTER-CLUSTER POWER AND DELAY CHARACTERISTICS

cluster index (NO.)	sub-6 GHz		mmWave	
	Relative delay (μs)	Relative power (dB)	Relative delay (μs)	Relative power (dB)
1	0.0977	−19.6694	0.0667	−14.6786
2	0.2336	−28.8640	0.1987	−25.8935
3	0.4371	−35.9763	0.4227	−32.7006
4	0.7062	−38.9410	0.8348	−34.1254
5	1.1425	−42.5525	1.3065	−34.9582
6	1.6719	−50.0298	1.5096	−37.6226
7	1.9996	−54.5792	1.6351	−40.3471
8	2.3094	−57.1457	2.2187	−64.0877
9	2.4600	−61.6080	—	—

C. Time-Varying Evolution Characterization

The high mobility of the UAV during flight leads to the fast time-varying characteristics of the channel. Therefore, the clusters undergo the B-D phenomenon. Fig. 13 shows the cluster tracking results for UAV flight time of 30 s. The new cluster index (CLID) is assigned to a new cluster, and the old cluster inherits the CLID of the previous cluster [28]. The number of whole clusters observed in the two bands differs throughout the time interval. Nearly 500 and 330 clusters are detected for sub-6 GHz and mmWave, respectively. The continuous distance on the x-axis indicates the life period of the relevant cluster. Many clusters exist only for a time period and there is clear B-D phenomenon. Also, many clusters exit only a stationary region and can not be tracked consistently. In addition, there is the LoS component that persists throughout flight as also shown in Fig. 13.

Fig. 14 shows the CDFs of the cluster lifetime for the UAV flight time of 30 s. This paper uses the log-normal distribution to analyze and model the lifetime. It can be found that the lifetime of the clusters that account for 80% of the total number of clusters for sub-6 GHz and mmWave are 1.0 s and 1.1 s, respectively. It illustrates the relatively short lifetime of the clusters during the whole UAV flight, especially in sub-6 GHz band.

Moreover, this paper introduces a Markov chain to well characterize the dynamic clusters evolution of UAV-to-ground communication channels. The state transition chart is shown in Fig. 15. Four typical states conform to the characteristics of UAV-to-ground communication channel evolution. By calculating the channel data, the state shift probability matrices for the specific sub-6 GHz T_1 and mmWave T_2 are shown as

$$T_1 = \begin{bmatrix} \xi_{00}^1 & \xi_{01}^1 & \xi_{02}^1 & \xi_{03}^1 \\ \xi_{10}^1 & \xi_{11}^1 & \xi_{12}^1 & \xi_{13}^1 \\ \xi_{20}^1 & \xi_{21}^1 & \xi_{22}^1 & \xi_{23}^1 \\ \xi_{30}^1 & \xi_{31}^1 & \xi_{32}^1 & \xi_{33}^1 \end{bmatrix} = \begin{bmatrix} 0.05 & 0.28 & 0.06 & 0.61 \\ 0.09 & 0.36 & 0.08 & 0.47 \\ 0.04 & 0.36 & 0.08 & 0.52 \\ 0.04 & 0.39 & 0.09 & 0.48 \end{bmatrix}$$

$$T_2 = \begin{bmatrix} \xi_{00}^2 & \xi_{01}^2 & \xi_{02}^2 & \xi_{03}^2 \\ \xi_{10}^2 & \xi_{11}^2 & \xi_{12}^2 & \xi_{13}^2 \\ \xi_{20}^2 & \xi_{21}^2 & \xi_{22}^2 & \xi_{23}^2 \\ \xi_{30}^2 & \xi_{31}^2 & \xi_{32}^2 & \xi_{33}^2 \end{bmatrix} = \begin{bmatrix} 0.29 & 0.47 & 0.11 & 0.13 \\ 0.28 & 0.49 & 0.09 & 0.14 \\ 0.20 & 0.59 & 0.07 & 0.14 \\ 0.23 & 0.44 & 0.12 & 0.21 \end{bmatrix},$$

where ξ_{ij}^1 and ξ_{ij}^2 represents the transfer probability from the i-th state to the j-th state for sub-6 GHz and mmWave, respectively. It can be observed that the probability of new clusters' generation and B-D phenomenon are higher regardless of which state they belonged to in the past, due to the large amount of scatterers and the rapid movement of the UAVs. This phenomenon also appears in fast-moving HSR communication scenarios [28].

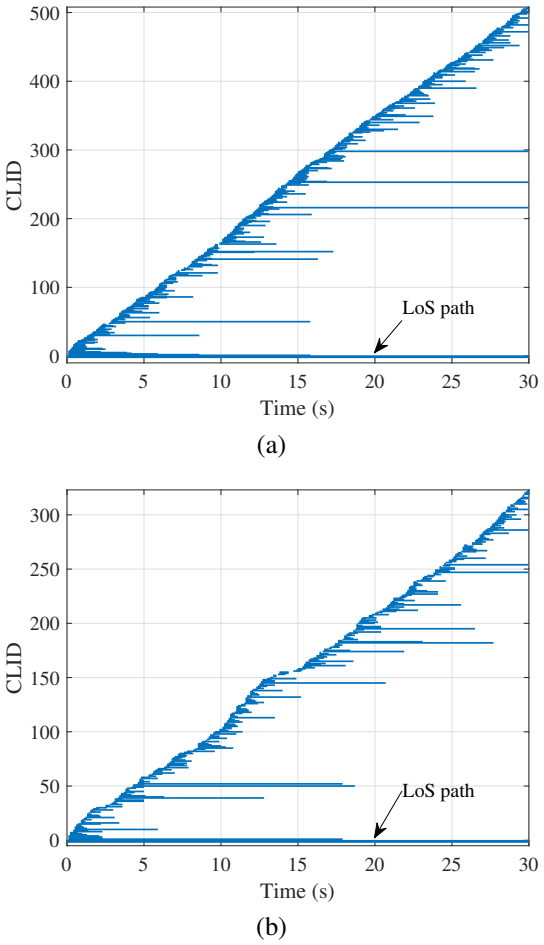


Fig. 13. Time-varying cluster tracking result in (a) sub-6 GHz band and (b) mmWave band.

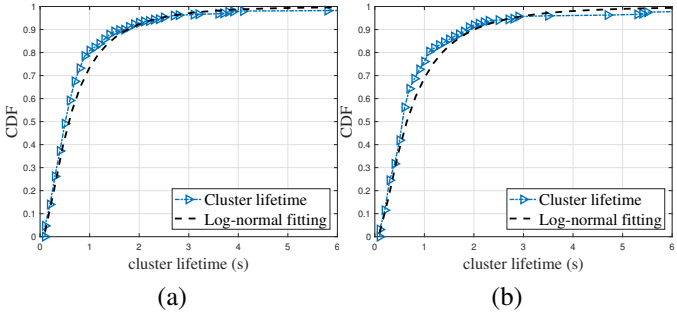


Fig. 14. The CDFs of cluster lifetime in (a) sub-6 GHz band and (b) mmWave band.

V. CLUSTER-BASED CHANNEL MODEL IMPLEMENTATION

In this section, a cluster-based UAV communications channel model is proposed based on the comprehensive channel characterization presented in Section IV. In the implementa-

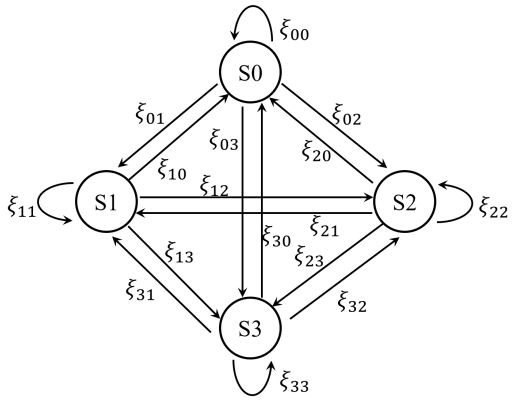


Fig. 15. Four-state Markov chain (S0: neither birth nor death; S1: only birth; S2: only death; S3: both birth and death) [28].

tion, the model parameters are used to obtain the LoS path, scatterer clusters, and MPCs within the clusters.

The specific model realization process is shown in Fig. 16. In the realization process, the initial step is to set the basic information, such as environmental parameters, frequency band, UAV flight trajectory, and 3-D distance between the Tx and Rx. Then, LoS paths and scatterer clusters need to be generated. The number of initial clusters is set to the maximum probability value based on the distribution presented in Fig. 12. Moreover, based on the inter-cluster characteristics, parameters such as delay and power of the clusters are generated by stochastic process [33]. According to the intra-cluster characteristics shown in Table III, the specific parameters of MPCs can be acquired through the corresponding distribution, such as the MPCs number within cluster, delay, power, etc. Once we have these parameters, we can derive the CIR by combining all the generated clusters and their corresponding intra-cluster MPCs [33]. The time-varying CIR is expressed as

$$h(t,\tau) = \sqrt{P_0(t)}e^{-j\phi_0(t)}\delta(\tau-\tau_0(t)) + \sum_{i=1}^{I(t)}\sum_{c=1}^{C_i(t)}\sqrt{P_{i,c}(t)}e^{-j\phi_{i,c}(t)}\delta(\tau-\tau_i(t)-\tau_{i,c}(t)) \tag{17}$$

where $I(t)$ is the number of clusters and $C_i(t)$ is the MPCs number in i-th cluster. The power and phase of the c-th ray in the i-th cluster are $P_{i,c}$ and $\phi_{i,c}$, respectively. The cluster delay and excess delay are τ_i and $\tau_{i,c}$. The Dirac delta function is $\delta(\cdot)$. Also specifically, the LoS ray parameters are set to P_0, ϕ_0, and τ_0, respectively. Note that the power, delay, and phase are both time-varying parameters with the flight.

In the evolutionary process, different clusters are given different lifetimes firstly, and the Markov chain generates the next state based on the current state. The state transfer diagram is shown in Fig. 15. Based on the state of generation, determine whether birth and death have occurred. When the Markov state is 0, it is assumed that the clusters have not changed. When the state is 1, we randomly generate new clusters and then generate intra-cluster MPCs based on the new clusters. The number of new clusters is generated using the Poisson process [41]. The lifetime of past clusters decreases with flight. The lifetime of

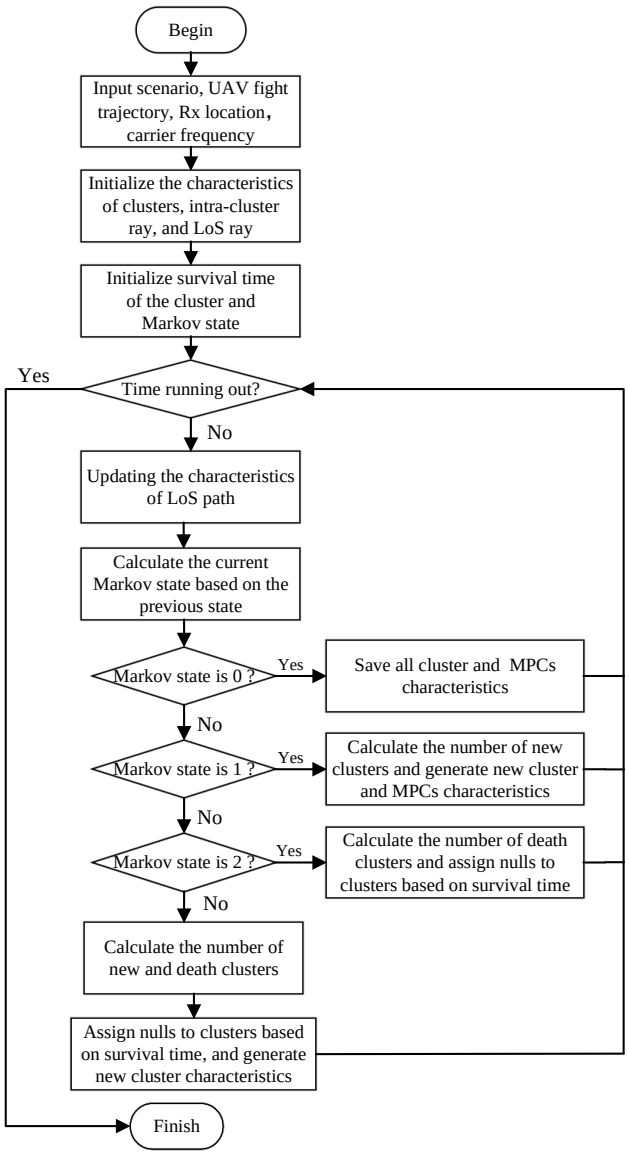


Fig. 16. The flowchart for the implementation of cluster-based B-D evolution channel model.

new clusters is randomly generated according to log-normal distribution. When the state is 2, some clusters are nulled based on the cluster's lifetime. When the state is 3, the B-D process of states 1 and 2 is accomplished. Therefore, simulation data of 10 s is generated based on the above steps.

Fig. 17 shows the delay PSDs of UAV-to-ground communications. It can be seen that the power at different periods is related to the delay at different time moments. The delay PSDs differ at t = 3 s and t = 5 s. The presence of significant LoS paths can be observed, and the distributions also differ between mmWave and sub-6 GHz bands. The number of MPCs may be concentrated at certain delays due to the presence of dense scatterers. Furthermore, there is considerable signal attenuation at mmWave band compared to sub-6 GHz band.

Fig. 18 compares the CDFs of RMS DS with the proposed model, measurement data, and reference model. RMS DS can be calculated from the square root of the second central mo-

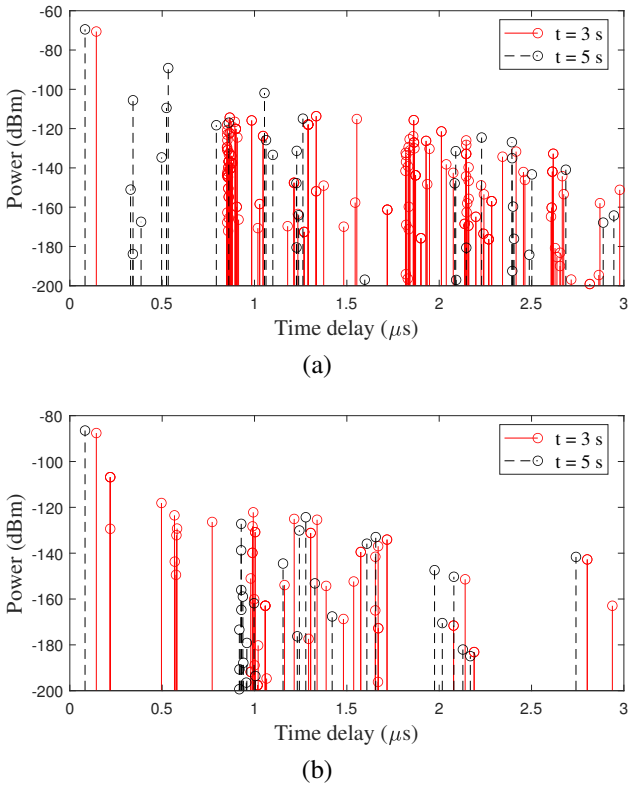


Fig. 17. The delay PSDs of the proposed UAV-to-ground channel model at t = 3 s and t = 5 s in (a) sub-6 GHz band and (b) mmWave band.

ment of the CIR generated by the stochastic process described above. It can be observed that the simulation and measurement channels exhibited good consistency. For 80% of the RMS DS, the difference between measurement and simulation data is 3 ns, which is relatively similar and acceptable. Note that the channel measurement data in Fig. 18 is from [42], and the frequency band selected here is 2.585 GHz, which has been used in current cellular networks. Besides, this data is obtained from the suburban scenario UAV-to-BS measurement environment. Due to the heavy weight of mmWave measurement equipment and the UAV's limited load, few mmWave UAV-to-ground communication channel measurements are available. Channel measurements at mmWave band will be carried out in the future. Moreover, our model is more consistent with the actual data compared to the WINNER+ suburban macrocells channel model [43]. Also, the RMS DS variance of the UAV-to-ground communication channels is smaller compared to the WINNER+ terrestrial channels. Overall, the proposed model can reflect the properties of the actual channel well. In addition, complex process operations and massive random parameters in large-scale GBSM are avoided. Therefore, this proposed model can provide a reference for UAV-to-ground communication network optimization and planning, and be used for channel simulator and emulator of dynamic UAV-to-ground communications.

VI. Conclusions

In this paper, a ML-based clustering and modeling method for UAV communication channels has been proposed. Compared with traditional methods, the proposed VB-GMM

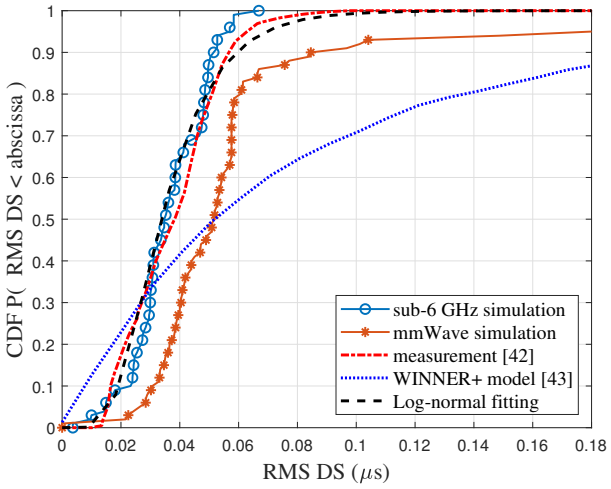


Fig. 18. The CDFs of RMS DS with simulation and measurement data.

method represents the actual channel situation well, and the number of clusters can be obtained automatically. Besides, the process of clustering and analyzing has considered space-time characterization and two typical frequency bands, i.e., sub-6 GHz and mmWave. The relationship between the intra-cluster MPCs and the fading slope has been analyzed. A comprehensive discussion of the number of MPCs, RMS DS/AS, delay/angle offset, and cluster power distributions has been conducted well. We also have analyzed the evolution of clusters during UAV flight. The log-normal distribution has been applied to fit the lifetime of the dynamic clusters, and clusters have frequent B-D phenomena during UAV flight. Based on the comprehensively analyzed intra-cluster, inter-cluster, and dynamic evolution characteristics, we have carried out the cluster-based UAV-to-ground communication channel modeling. A four-state Markov chain has been introduced to portray the B-D process in detail. Besides, channel measurement has been well-fitted to the simulation data. Therefore, these results can provide accurate and low-complexity support for the design and evaluation of future 6G UAV communication systems.

References

[1] C.-X. Wang, J. Huang, H. Wang, X. Gao, X. You, and Y. Hao, "6G wireless channel measurements and models: Trends and challenges," *IEEE Veh. Technol. Mag.*, vol. 15, no. 4, pp. 22–32, 2020.

[2] B. Li, Z. Fei, and Y. Zhang, "UAV communications for 5G and beyond: Recent advances and future trends," *IEEE Int. Things J.*, vol. 6, no. 2, pp. 2241–2263, 2018.

[3] X. Cheng, Z. Huang, and L. Bai, "Channel nonstationarity and consistency for beyond 5G and 6G: A survey," *IEEE Commun. Surveys & Tutorials*, vol. 24, no. 3, pp. 1634–1669, 2022.

[4] S. Yang, Z. Zhang, J. Zhang, and J. Zhang, "Impact of rotary-wing UAV wobbling on millimeter-wave air-to-ground wireless channel," *IEEE Trans. Veh. Technol.*, vol. 71, no. 9, pp. 9174–9185, 2022.

[5] C.-X. Wang, Z. Lv, X. Gao, X. You, Y. Hao, and H. Haas, "Pervasive wireless channel modeling theory and applications to 6G GBSMs for all frequency bands and all scenarios," *IEEE Trans. Veh. Technol.*, vol. 71, no. 9, pp. 9159–9173, 2022.

[6] G. Sun, R. He, B. Ai, C. Huang, and Z. Zhong, "Dynamic clustering of multipath components for time-varying propagation channels," *IEEE Trans. Veh. Technol.*, vol. 70, no. 12, pp. 13 396–13 400, 2021.

[7] L. Bai, Z. Huang, and X. Cheng, "A non-stationary model with time-space consistency for 6G massive MIMO mmWave UAV channels," *IEEE Trans. Wireless Commun.*, vol. 22, no. 3, pp. 2048–2064, 2023.

[8] H. Chang, C.-X. Wang, J. Bian, R. Feng, Y. He, Y. Chen, and E.-H. M. Aggoune, "A novel 3D beam domain channel model for UAV massive MIMO communications," *IEEE Trans. Wireless Commun.*, vol. 22, no. 8, pp. 5431–5445, 2023.
[9] Y. Liu, C.-X. Wang, H. Chang, Y. He, and J. Bian, "A novel non-stationary 6G UAV channel model for maritime communications," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 10, pp. 2992–3005, 2021.
[10] Q. Zhu *et al.*, "Map-based channel modeling and generation for U2V mmWave communication," *IEEE Trans. Veh. Technol.*, vol. 71, no. 8, pp. 8004–8015, 2022.
[11] H. Jiang, Z. Zhang, and G. Gui, "Three-dimensional non-stationary wideband geometry-based UAV channel model for A2G communication environments," *IEEE Access*, vol. 7, pp. 26 116–26 122, 2019.
[12] Y. Liu, C.-X. Wang, J. Huang, J. Sun, and W. Zhang, "Novel 3-D nonstationary mmWave massive MIMO channel models for 5G high-speed train wireless communications," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2077–2086, 2018.
[13] I. T. Union, "IMT-2020 Preliminary Draft New Report ITU-R M.[IMT-2020.EVAL]," Tech. Rep., June 2017.
[14] 5GCMSIG, "5G channel model for bands up to 100 GHz, v2.0," Tech. Rep., Mar. 2016.
[15] 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz, V16.1.0," Tech. Rep., Jan. 2020.
[16] H. Jiang *et al.*, "A novel 3D UAV channel model for A2G communication environments using AoD and AoA estimation algorithms," *IEEE Trans. Wireless Commun.*, vol. 68, no. 11, pp. 7232–7246, 2020.
[17] Y. Yang *et al.*, "6G network AI architecture for everyone-centric customized services," *IEEE Network*, vol. 37, no. 5, pp. 71–80, 2023.
[18] A. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE J. Sel. Areas Commun.*, vol. 5, no. 2, pp. 128–137, 1987.
[19] T. Jiang, J. Zhang, M. Shafi, L. Tian, and P. Tang, "The comparative study of S-V model between 3.5 and 28 GHz in indoor and outdoor scenarios," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 2351–2364, 2020.
[20] C. Gustafson, K. Haneda, S. Wyne, and F. Tufvesson, "On mm-Wave Multipath Clustering and Channel Modeling," *IEEE Trans. Antennas Propag.*, vol. 62, no. 3, pp. 1445–1455, 2014.
[21] J. Lee, "Cluster-based millimeter-wave outdoor-to-indoor propagation characteristics based on 32 GHz measurement analysis," *IEEE Antennas and Wireless Propaga. Lett.*, vol. 20, no. 1, pp. 73–77, 2021.
[22] J. Qiu, C. Tao, L. Liu, Z. Lin, and Z. Tan, "A novel hybrid CDL-based multipath propagation model for the high-speed railway at 2.35 GHz," *Chin. Sci. Bull.*, vol. 59, no. 35, pp. 4976–4987, 2014.
[23] Y. Yang, J. Sun, W. Zhang, C.-X. Wang, and X. Ge, "Ray tracing based 60 GHz channel clustering and analysis in staircase environment," in *Proc. GLOBECOM*, 2017, pp. 1–5.
[24] C. Gentile, "Using the Kurtosis measure to identify clusters in wireless channel impulse responses," *IEEE Trans. Antennas Propag.*, vol. 61, no. 6, pp. 3392–3395, 2013.
[25] C. Huang *et al.*, "A power-angle-spectrum based clustering and tracking algorithm for time-varying radio channels," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 291–305, 2019.
[26] X. Cai, B. Peng, X. Yin, and A. P. Yuste, "Hough-transform-based cluster identification and modeling for V2V channels based on measurements," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 3838–3852, 2018.
[27] Y. Li, J. Zhang, Z. Ma, and Y. Zhang, "Clustering analysis in the wireless propagation channel with a variational Gaussian mixture model," *IEEE Trans. Big Data*, vol. 6, no. 2, pp. 223–232, 2020.
[28] T. Zhou, Y. Qiao, S. Salous, L. Liu, and C. Tao, "Machine learning-based multipath components clustering and cluster characteristics analysis in high-speed railway scenarios," *IEEE Trans. Antennas Propag.*, vol. 70, no. 6, pp. 4027–4039, 2022.
[29] K. Mahler, W. Keusgen, F. Tufvesson, T. Zemen, and G. Caire, "Tracking of wideband multipath components in a vehicular communication scenario," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 15–25, 2017.
[30] M. Gan, Z. Xu, C. F. Mecklenbräuker, and T. Zemen, "Cluster lifetime characterization for vehicular communication channels," in *Proc. EuCAP*, 2015, pp. 1–5.
[31] P. Hanpinitsak, K. Saito, W. Fan, J. Hejselbæk, J.-I. Takada, and G. F. Pedersen, "Frequency characteristics of geometry-based clusters in indoor hall environment at SHF bands," *IEEE Access*, vol. 7, pp. 75 420–75 433, 2019.
[32] Q. Wang *et al.*, "A framework of automatic clustering and tracking for time-variant multipath components," *IEEE Commun. Lett.*, vol. 21, no. 4, pp. 953–956, 2017.
[33] Z. Cui, K. Guan, C. Oestges, C. Briso-Rodríguez, B. Ai, and Z. Zhong, "Cluster-based characterization and modeling for UAV air-to-ground time-varying channels," *IEEE Trans. Veh. Technol.*, vol. 71, no. 7, pp. 6872–6883, 2022.
[34] X. Cai *et al.*, "Empirical low-altitude air-to-ground spatial channel characterization for cellular networks connectivity," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 10, pp. 2975–2991, 2021.
[35] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning*. Springer, 2006.
[36] H. Schütze, C. D. Manning, and P. Raghavan, *Introduction to information retrieval*. Cambridge University Press Cambridge, 2008.
[37] Y. Li, Y. Chen, D. Yan, K. Guan, and C. Han, "Channel characterization and ray-tracing assisted stochastic modeling for urban vehicle-to-infrastructure Terahertz communications," *IEEE Trans. Veh. Technol.*, vol. 72, no. 3, pp. 2748–2763, 2023.
[38] Remcom, "Wireless InSite." https://www.remcom.com/wireless-insite-em-propagation-software.
[39] ITU-R, "P. 2040-2. Effects of building materials and structures on radiowave propagation above about 100 MHz," *ITU Recommendations, Tech. Rep*, 2021.
[40] M. Yang *et al.*, "Measurements and cluster-based modeling of vehicle-to-vehicle channels with large vehicle obstructions," *IEEE Trans. Wireless Commun.*, vol. 19, no. 9, pp. 5860–5874, 2020.
[41] H. Chang *et al.*, "A novel nonstationary 6G UAV-to-ground wireless channel model with 3-D arbitrary trajectory changes," *IEEE Int. Things J.*, vol. 8, no. 12, pp. 9865–9877, 2020.
[42] X. Cai *et al.*, "An empirical air-to-ground channel model based on passive measurements in LTE," *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1140–1154, 2018.
[43] J. Meinilä *et al.*, "WINNER+ final channel models," *docment CELTIC/CP5-026 D5.3*, 2010.

Zhaolei Zhang received the B.E. degree in electronic information engineering from Shandong University of Science and Technology, Qingdao, China, in 2022. He is currently pursuing his master's degree with the School of Integrated Circuits, Shandong University, Jinan, China. His current research interests include channel measurements and modeling, unmanned aerial vehicle (UAV) communications, and artificial intelligence in wireless communications.

Yu Liu (Member, IEEE) received the Ph.D. degree in communication and information systems from Shandong University, Jinan, China, in 2017. From 2015 to 2017, she was a visiting scholar with the School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, U.K.. From 2017 to 2019, she was a Postdoctoral Research Associate with the School of Information Science and Engineering, Shandong University, Jinan, China. Since 2019, she has been an Associate Professor with the School of Integrated Circuits, Shandong University, Jinan, China. Her main research interests include nonstationary wireless MIMO channel modeling, high mobility channel characterization and modeling, and AI based channel modeling.

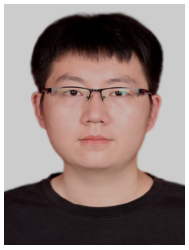
Cheng-Xiang Wang (Fellow, IEEE) received the B.Sc. and M.Eng. degrees in communication and information systems from Shandong University, China, in 1997 and 2000, respectively, and the Ph.D. degree in wireless communications from Aalborg University, Denmark, in 2004.

He was a Research Assistant with the Hamburg University of Technology, Hamburg, Germany, from 2000 to 2001, a Visiting Researcher with Siemens AG Mobile Phones, Munich, Germany, in 2004, and a Research Fellow with the University of Agder, Grimstad, Norway, from 2001 to 2005. He was with Heriot-Watt University, Edinburgh, U.K., from 2005 to 2018, where he was promoted to a professor in 2011. He has been with Southeast University, Nanjing, China, as a professor since 2018, and he is now the Executive Dean of the School of Information Science and Engineering. He is also a professor with Pervasive Communication Research Center, Purple Mountain Laboratories, Nanjing, China. He has authored 4 books, 3 book chapters, and over 550 papers in refereed journals and conference proceedings, including 27 highly cited papers. He has also delivered 27 invited keynote speeches/talks and 18 tutorials in international conferences. His current research interests include wireless channel measurements and modeling, 6G wireless communication networks, and electromagnetic information theory.

Dr. Wang is a Member of the Academia Europaea (The Academy of Europe), a Member of the European Academy of Sciences and Arts (EASA), a Fellow of the Royal Society of Edinburgh (FRSE), IEEE, IET and China Institute of Communications (CIC), an IEEE Communications Society Distinguished Lecturer in 2019 and 2020, a Highly-Cited Researcher recognized by Clarivate Analytics in 2017-2020. He is currently an Executive Editorial Committee Member of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He has served as an Editor for over ten international journals, including the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, from 2007 to 2009, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, from 2011 to 2017, and the IEEE TRANSACTIONS ON COMMUNICATIONS, from 2015 to 2017. He was a Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, Special Issue on Vehicular Communications and Networks (Lead Guest Editor), Special Issue on Spectrum and Energy Efficient Design of Wireless Communication Networks, and Special Issue on Airborne Communication Networks. He was also a Guest Editor for the IEEE TRANSACTIONS ON BIG DATA, Special Issue on Wireless Big Data, and is a Guest Editor for the IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING, Special Issue on Intelligent Resource Management for 5G and Beyond. He has served as a TPC Member, a TPC Chair, and a General Chair for more than 30 international conferences. He received 16 Best Paper Awards from IEEE GLOBECOM 2010, IEEE ICCT 2011, ITST 2012, IEEE VTC 2013 Spring, IWCMC 2015, IWCMC 2016, IEEE/CIC ICCC 2016, WPMC 2016, WOCC 2019, IWCMC 2020, WCSP 2020, CSPS2021, WCSP 2021, IEEE/CIC ICCC 2022, and IEEE ICCT 2023.

Hengtai Chang (Member, IEEE) received the B.Sc. and Ph.D. degrees from the School of Information Science and Engineering, Shandong University, China, in 2016 and 2021, respectively. He is currently a Post-Doctoral Research Associate at Purple Mountain Laboratories, China, and a Post-Doctoral Research Associate at the China National Mobile Communications Research Laboratory, Southeast University, China. His current research interests include UAV communications, wireless propagation channel measurements and channel modeling, and B5G/6G wireless communications.

Ji Bian (Member, IEEE) received the B.Sc. degree in electronic information science and technology from Shandong Normal University, Jinan, China, in 2010, the M.Sc. degree in signal and information processing from Nanjing University of Posts and Telecommunications, Nanjing, China, in 2013, and the Ph.D. degree in information and communication engineering from Shandong University, Jinan, China, in 2019. From 2017 to 2018, he was a visiting scholar with the School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, U.K. He is currently an Associate Professor with the School of Information Science and Engineering, Shandong Normal University, Jinan, China. His research interests include 6G channel modeling and wireless big data.



Jingfan Zhang received the B.E. degree in School of Electronic Information Engineering, Shandong University of Science and Technology, Qingdao, China, in 2020. He is currently pursuing his master's degree with School of Integrated Circuits, Shandong University, Jinan, China. His current research interests include UAV communications and wireless propagation channel measurements and modeling.